PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/11, C07K 14/00

A2

(11) International Publication Number: WO 96/13583

(43) International Publication Date: 9 May 1996 (09.05.96)

(21) International Application Number: PCT/EP95/04117

(22) International Filing Date: 20 October 1995 (20.10.95)

(30) Priority Data:
94116558.1 20 October 1994 (20.10.94) EP

(34) Countries for which the regional or international application was filed: DE et al.

(71) Applicant *(for all designated States except US)*: MORPHOSYS GESELLSCHAFT FÜR PROTEINOPTIMIERUNG MBH [DE/DE]; Frankfurter Ring 193a, D-80807 Munich (DE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PACK, Peter [DE/DE]; Franz-Wolter-Strasse 4/III, D-81925 Oberföhring (DE). LUPAS, Andrei [DE/DE]; Fasanenstrasse 68a, D-82008 Unterhaching (DE).

(74) Agent: VOSSIUS & PARTNER; P.O. Box 86 07 67, D-81634 Munich (DE).

(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES

(57) Abstract

The present invention relates to a method for targeted assembly of distinct active peptide or protein domains into a single complex and to such complexes. The invention relates particularly to the fusion of peptide or protein domains to complementary association domains which are derived from a single tertiary or quaternary structure by segmentation. The association domains are designed to assemble in a complementary fashion, thereby providing multifunctional (poly)peptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgyzstan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES

Background of the Invention

Increasingly, there is a need for proteins which combine two or more functions, such as binding or catalysis, in a single structure. Typically, proteins which combine two or more functions are prepared either as fusion proteins or through chemical conjugation of the component functional domains. Both of these approaches suffer from disadvantages. Genetic "single chain" fusions suffer the disadvantages that (i) only a few (2-3) proteins can be fused (Rock et al., 1992, *Prot. Eng. 5*, 583-591), (ii) mutual interference between the component domains may hinder folding, and (iii) the size of the fusion protein may make it difficult to prepare. The alternative, chemical cross-linking *in vitro* following purification of independently expressed proteins, is difficult to control and invariably leads to undefined products and to a severe loss in yield of functional material.

Recently, methods for achieving non-covalent association of two or more of the same functional domains have been developed. This can be achieved through the use of domains attached to peptides which self-associate to form homo-multimers (Pack & Plückthun, 1992, *Biochemistry 31*, 1579-1584). For example, the association of two separately expressed scFv antibody fragments by C-terminally fused amphipathic helices *in vivo* provides homo-dimers of antibody fragments in *E. coli* (PCT/EP93/00082; Pack et al., 1993, *Bio/Technology 11*, 1271-1277) or homo-tetramers (Pack et al., 1995, *J. Mol. Biol., 246, 28-34*).

To assemble distinct protein functions such as two antibody fragments with different specificities fused to such association domains, the helices must have a tendency to form hetero-multimers. In principle, this could be achieved with complementary helices such as the hetero-dimerizing JUN and FOS zippers of the AP-1 transcription factor (O'Shea et al., 1992, *Cell 68*, 699-708). The clear disadvantage of association domains based on hetero-associating helices,

however, is their pseudo-symmetry and their similar periodicity of hydrophobic and hydrophilic residues. This structural similarity results in a strong tendency to form homo-dimers and, thus, to lower significantly the yield of hetero-dimers (O'Shea et al., 1992, *Cell 68*, 699-708; Pack, 1994, Ph. D. thesis, Ludwig-Maximilians-Universität München). Furthermore, the formation of JUN/FOS hetero-dimers is kinetically disfavoured and requires a temperature-dependent unfolding of the kinetically favoured homo-dimers, especially JUN/JUN homo-dimers (PCT/EP93/00082; O'Shea et al., 1992, *Cell 68*, 699-708; Pack, 1994, Ph. D. thesis, Ludwig-Maximilians-Universität München). Because of the need for additional purification steps to separate the unwanted homo-dimers from hetero-dimers and the resulting decrease in yield, hetero-association domains based on amphipathic helices do not result in practical advantages compared to conventional chemical coupling.

These disadvantages of the prior art are overcome by the present invention which provides multi-functional polypeptides and methods for the preparation of these multi-functional proteins. This is achieved via the use of association domains which are designed to associate predominantly in a complementary fashion, and not to self-associate.

Detailed Description of the Invention

In the earliest steps of protein folding, peptide chains form a disordered hydrophobic core by collapsing hydrophobic residues into the interior of an intermediate "molten globule". This hydrophobic effect is considered to be the most important driving force of folding (Matthews, 1993, *Annu. Rev. Biochem. 62*, 653 - 683; Fersht, 1993, *FEBS Letters 325*, 5 - 16). The burial of hydrophobic residues and the resulting exclusion of solvent is the determining factor in the stability of compact tertiary structures such as acyl-phosphatase (Pastore et al, J. Mol. Biol. 224, 427-440, 1992) interleukin-2 (Brandhuber et al., 1987, *Science 238*, 1707 - 1709), calbindin (Parmentier, 1990, *Adv. Exp. Med. Biol. 269*, 27-34) or ubiquitin (Briggs & Roder, 1992, *Proc. Natl. Acad. Sci. USA 89*, 2017 - 2021).

This concept forms the basis of the present invention, which provides individually encoded peptides or "segments" which, in a single continuous chain, would comprise a compact tertiary structure with a highly hydrophobic core. The component peptides are chosen so as to be asymmetric in their assumed structure, so as not to self-associate to form homo-multimers, but rather to associate in a complementary fashion, adopting a stable complex which resembles the parent tertiary structure. On the genetic level, these segments are encoded by interchangeable cassettes with suitable restriction sites. These standardized cassettes are fused C- or N-terminally to different recombinant proteins via a linker or hinge in a suitable expression vector system.

Thus, the present invention relates to a multi-functional polypeptide comprising:

(a) a first amino acid sequence attached to at least one functional domain;

(b) a second amino acid sequence attached to at least one further functional domain; and

(c) optionally, further amino acid sequences each attached to at least one further functional domain;

wherein any one or more of said amino acid sequences interacts with at least one of said amino acid sequences in a complementary fashion to form a parental, native-like tertiary or optionally quaternary structure and wherein the parental, native-like tertiary or optionally quaternary structure is derived from a single parent polypeptide. In this context, the term parent polypeptide refers to a polypeptide which has a compact tertiary or quarternary structure with a hydrophobic core. The invention provides for many different parent polypeptides to be used as the basis for the association domain. Suitable polypeptides can be identified by searching for compact, single-domain proteins or protein fragments in the database of known protein structures (Protein Data Bank, PDB) and selecting structures that are stable and can be expressed at high yields in recombinant form. These structures can then be analyzed for hydrophobic sub-clusters by the method of Karpeisky and Ilyn (1992, *J. Mol. Biol. 224*, 629-638) or for structural units (such as ß-elements or helical hairpin structures) by standard molecular modelling techniques. In a further embodiment, the present invention

provides for multi-functional polypeptides wherein the single parent polypeptide is taken from the list ubiquitin, acyl phosphatase, IL-2, calbindin and myoglobin.

In a preferred embodiment, the present invention provides a multi-functional polypeptide comprising two or more amino acid sequences each attached to at least one functional domain, wherein any two or more of said amino acid sequences can associate in a complementary fashion to provide a parental, native like, tertiary or optionally quaternary structure.

Once structural sub-domains are identified, the protein is dissected in such a way these sub-domains remain intact. The selection process can be expanded to proteins for which no structure is available but which satisfy the criteria of stability and good expression. For these proteins, folding sub-domains can be determined by hydrogen exchange pulse-labelling of backbone amides during the folding reaction, followed by NMR detection in the native state (Roder et al., 1988, *Nature 355*, 700-704; Udgaonkar & Baldwin, 1988, *Science 255*, 594-597). Alternatively, folding sub-domains can be identified by mild proteolysis, denaturation, purification of fragments and reconstitution *in vitro* (Tasayco & Carey, 1992, *Science 255*, 594-597; Wu et al., 1993, *Biochemistry 32*, 10271-10276). Finally, additional clues for the choice of cleavage sites can be obtained from the exon structure in the case of eukaryotic proteins, since the exons frequently (though not always) correspond to structural sub-domains of a protein. This has, for example, been discussed for the case of myoglobin (Go 1981, *Nature 291*, 90).

The yield of properly assembled molecules is expected to decrease significantly for constructs in which a protein domain is divided into three or more parts. This is due to the fact that several sub-domains must come together simultaneously to form a viable structure. This effect is countered by dividing the polypeptide chain into sub-domains that represent folding units (identified by the methods described above). Thus, not only the final, assembled complex but also assembly intermediates will have the stability necessary to allow their accumulation in the

host during expression, resulting in a greatly improved kinetic behaviour of the system.

In solution, the isolated segments have little secondary structure and remain monomeric or form transient, non-specific and easily disrupted aggregates. Only upon mixing, either by separate expression and purification, or by co-expression, can the concerted folding of complementary segments provide the necessary intermediate interaction of residues (Matthews, 1993, *Annu. Rev. Biochem. 62*, 653 - 683) that results in the formation of a compact, native-like structure. This association, mainly driven by the burial of hydrophobic residues of all segments into a single hydrophobic core, leads to a targeted assembly of the N- or C-terminally fused proteins to a multi-functional complex *in vivo* or *in vitro*.

Optionally, the reconstituted native-like structure may also contribute an enzymatic or binding activity to increase the number of effector functions in the assembled complex. Accordingly, the present invention also provides a multi-functional polypeptide as described above, in which the native-like, tertiary or quaternary structure provides a biological activity. For example, when acyl phosphatase is used as the basis of the association domain, it is expected that the multi-functional polypeptide will retain some phosphatase activity.

The present invention provides for many different types of functional domains to be linked into the multi-functional polypeptide. Particularly preferred are cases in which one or more, preferably two, of said functional domains are fragments derived from molecules of the immunoglobulin superfamily. In particularly preferred embodiments, said fragments are antibody fragments. Also preferred are cases in which at least one of the functional domains possesses biological activity other than that associated with a fragment derived from a member of the immunoglobulin superfamily. By way of example, the present invention provides for the targeted assembly of enzymes, toxins, cytokines, peptide hormones, immunoglobulins, metal binding domains, soluble receptors, lectins, lipoproteins, purification tails and bioactive peptides to multi-functional complexes (Fig. 1) based on a modular system of expression vectors, restriction sites and "plug-in" gene cassettes coding for assembly segments, peptide linkers and functional domains (Fig.2).

If covalent linkage between the segments is necessary to prevent dissociation at low concentrations, cysteines can be introduced to form inter-segmental disulphide bridges between the amino acid sequences which comprise the association domain (Ecker et al., 1989,*J. Biol. Chem. 264*, 1887-1893; Pack & Plückthun, 1992, *Biochemistry 31*, 1579-1584). Accordingly, the present invention provides multi-functional polypeptides wherein the folding of the component amino acid sequences is stabilized by a covalent bond.

In order to provide some flexibility between the association domain and the appended functional domains, it may be desired to incorporate a linker peptide. Accordingly, the present invention provides for multi-functional polypeptides of the type described above wherein at least one of the functional domains is coupled to said amino acid sequence via a flexible peptide linker. By way of example, the flexible linker may be derived from the hinge region of an antibody.

The invention enables even more complex multi-functional polypeptides to be constructed via the attachment of at least one further (poly)peptide to one or more of said amino acid sequences. By way of example, the further (poly)peptide can be taken from the list enzymes, toxins, cytokines, peptide hormones, immunoglobulins, metal binding domains, soluble receptors, lectins, lipoproteins, purification tails, in particular peptides which are able to bind to an independent binding entity, bioactive peptides, preferably of 5 to 15 amino acid residues, metal binding proteins, DNA binding domains, transcription factors and growth factors.

For therapeutic purposes, it is often desirable that proteinaceous substances display the minimum possible immunogenicity. Accordingly, the present invention provides for multi-functional polypeptides as described above in which at least one of said amino acid sequences, functional domains, or further (poly)peptides is of human origin.

In addition to the peptides and proteins provided above, the present invention also provides for DNA sequences, vectors, preferably bicistronic vectors, vector cassettes, characterised in that they comprise a DNA sequence encoding an amino acid sequence and optionally at least one further (poly)peptide comprised in the multifunctional polypeptide of the invention, and additionally at least one,

preferably singular cloning sites for inserting the DNA encoding at least one further functional domain or that they comprise DNA sequences encoding the amino acid sequences, and optionally the further (poly)peptide(s) comprised in the multifunctional polypeptide of the invention and suitable restriction sites for the cloning of DNA sequences encoding the functional domains, such that upon expression of the DNA sequences after the insertion of the DNA sequences encoding the functional domains into said restriction sites, in a suitable host the multifunctional polypeptide of the invention is formed. In a preferred embodiment said vector cassette is characterised in that it comprises the inserted DNA sequence(s) encoding said functional domain(s) and host cells transformed with at least one vector or vector cassette of the invention which can be used for the preparation of said multi-functional polypeptides.

In a further preferred embodiment, said host cell is a mammalian, preferably human, yeast, insect, plant or bacterial, preferably E. coli cell.

The invention further provides for a method for the production of a multifunctional polypeptide of the invention, which comprises culturing the host cell of the invention in a suitable medium, and recovering said multifunctional polypeptide produced by said host cell.

In a further embodiment, the invention relates to a method for the production of a multifunctional polypeptide of the invention which comprises culturing at least two host cells of the invention in a suitable medium, said host cells each producing only one of said first and said second amino acid sequences attached to at least one further functional domain, recovering the amino acid sequences, mixing thereof under mildly denaturing conditions and allowing in vitro folding of the multifunctional polypeptide of the invention from said amino acid sequences.

In a particular preferred embodiment, said method is characterised in that the further amino acid sequences attached to at least one further functional domain

are/is produced by at least one further host cell not producing said first or second amino acid sequence.

In another particularly preferred embodiment of the invention, said method is characterised in that at least one further amino acid sequence attached to at least one further functional domain is produced by the host cell of the invention producing said first or second amino acid sequence.

In further preferred embodiments, the present invention provides for pharmaceutical and diagnostic compositions comprising the multi-functional polypeptides described above, said pharmaceutical compositions optionally comprising a pharmaceutically acceptable carrier. Finally, the invention provides for a kit comprising one or more vector cassettes useful in the preparation of said multi-functional polypeptides.

The invention is now illustrated by reference to the following examples, which are provided for the purposes of illustration only and are not intended to limit the scope of the invention.

Example 1: Segmented human ubiquitin as an assembly device

Ubiquitin is a compact intracellular protein of only 76 residues (Fig. 3) and a molecular weight of 5 kDa. It shows the highest conservation among all known proteins and is involved in the degradation pathway of intracellular eukaryotic proteins by forming intermediate isopeptide bonds to its C-terminus and to Lys48 (Hershko & Ciechanover, 1992, *Ann. Rev. Biochem. 61,* 761-807).

To use ubiquitin as an assembly device, the unwanted function can be abolished by truncation of the last three C-terminal residues (--Arg-Gly-Gly), and the exchange of Lys48 to Arg, which prevents the formation of isopeptide bonds to this residue. The altered sequence is then divided in a loop at position Gly36, so that the hydrophobic core falls apart into two segments (called ALPHA and BETA). The synthetic nucleotide sequence of the segments (Fig 4, 5) carry appropriate restriction sites (MroI-HindIII) at the termini, so that the cassette encoding the segments can be easily ligated to a EcoRI-MroI cassette encoding the flexible linker (hinge of hulgG3; Fig. 6). The cassettes are inserted into the expression vector pIG3 (EcoRI-HindIII; Fig. 7) encoding the scFv fragment of the antibody McPC603 under the lac promoter/operator (Ge et al., 1995, in: *Antibody engineering: A practical approach.* IRL Press, New York, Borrebaeck ed., 229-261). Insertion of a second functional fragment (scFv fragment of the anti-ß-lactam antibody 2H10 with phoA signal sequence) linked to association segment BETA as an XbaI-HindIII DNA fragment (Fig. 8) results in a di-cistronic expression vector (Pack, 1994, Ph. D. thesis, Ludwig-Maximilians-Universität München). After induction with IPTG and translation, the signal sequences guide the antibody fragments fused to the assembly segments to the periplasm, where they assemble to a complex with a reconstituted native-like ubiquitin fold and two different antibody specificities. The complex, a bispecific immunoglobulin, can be recovered and purified by affinity chromatography of cell extract (Pack, 1994, Ph. D. thesis, Ludwig-Maximilians-Universität München).

Example 2: Covalent linkage of the native-like tertiary structure of the assembly device by engineered disulphide bridges and combination of a C-terminal peptide linker with an in-frame restriction site.

The conformational stability of undivided, native ubiquitin can be enhanced by introduction of disulfides at positions 4 and 66 without perturbation in the backbone (Ecker et al., 1989, *J. Biol. Chem. 264*, 1887-1893; Fig. 9). In the context of this invention, the engineering of disulfide bridges provides the covalent linkage of segments (Fig. 10, 11) after co-folding and assembly.
To raise the number of possible functional domains in the assembled complex, a C-terminal peptide can be fused to one or more of the segments of the assembly device. To fuse a functional domain like an enzyme, cytokine, antibody fragment, purification peptide or toxin to this linker, a restriction site, preferably unique, has to be introduced in-frame (Fig. 11). Gene synthesis, cloning, expression as well as recovery of the assembled, covalently linked complex is according to example 1.

Example 3: Segmented human interleukin-2 (IL2) as an assembly device

Human interleukin-2 (Brandhuber et al., 1987, *Science 238*, 1707 - 1709; Kuziel & Greene, 1991, in: *The Cytokine Handbook. Academic Press,* 84-100) is used as an assembly device by segmentation between position His79 and Lys 80 (Fig. 12). The device, encoded by MroI-AscI-HindII gene cassettes (Fig. 13, 14) combines the low immunogenicity of the plasmatic protein with a preferable effector function of the native-like cytokine structure and an inter-segmental cysteine bridge (Cys58-Cys105) after assembly. The combination of one or more antibody fragments against tumor antigens with additional cytokines like IL6 or

IL12 targets the multi-cytokine complex (Rock et al., 1992, *Prot. Eng. 5*, 583-591) directly to the tumour.

Example 4 : Segmented human apomyoglobin as an assembly device with three segments

To use more than two segments of a native structure as an assembly device, the hydrophobic interface between the segments has to be large enough to provide the sufficient hydrophobic interaction for non-covalent linkage. Myoglobin (Fig.15) is expressible in large amounts in *E. coli* (Guillemette et al., 1991, *Protein Eng. 4*, 585-592). Up to six functional domains can be assembled by a threefold segmented structure (Fig. 16, 17, 18), three at the N-termini and three at the C-termini of the segments. The presence of heme additionally stabilizes the native-like apomyoglobin fold and can be used as a switch to influence the association constant of the multi-functional complex.

Example 5: Bioactive peptides as functional domains

Certain peptides derived from amphipathic loop structures of LPS-binding proteins (Hoess et al., 1993, *EMBO J. 12*, 3351-3356) are able to neutralize endotoxin. This effect is enhanced by multivalent display of these short peptides (10-15 residues; Hoess, unpublished results). The present invention provides a method to express and assemble several of short peptides (Fig.19), fused to an assembly segment, in a multivalent complex or in combination with other functional domains. The peptides can be fused either to the N-or to the C-terminus (Fig. 20, 21) of the assembly domain via the peptide linkers.

Example 6: A purification tail for IMAC as a functional domain

Peptide tails consisting of histidines are able to coordinate metal ions. They are used for purification of native proteins in immobilized metal affinity chromatography (IMAC). Multivalent display of the purification tail considerably improves the maximum purity achievable by IMAC (Lindner et al., 1992, *Methods: a companion to methods in enzymology 4*, 41-56). One or more gene cassettes (Fig. 22) encoding a polyhistidine tail can be fused to the assembly segment to provide a simple and efficient purification method for multi-functional complexes.

Example 7: The platelet aggregation inhibitor decorsin as a functional domain

Decorsin, a 39 residue protein of the leech *Macrobdella decora* (Fig. 23), acts as a potent antagonist of the platelet glycoprotein IIb-IIIa (Seymour et al., 1990, *J. Biol. Chem. 265*, 10143-10147). The gene cassette encoding the decorsin can be fused C- or N-terminally to an association segment (Fig. 24, 25). In arterial thrombotic deseases, a multivalent decorsin complex combined with an anti-fibrin antibody fragment can act as a powerful antithrombotic agent.

Claims

1. A multifunctional polypeptide comprising:
   (a) a first amino acid sequence attached to at least one functional domain;
   (b) a second amino acid sequence attached to at least one further functional domain; and
   (c) optionally, further amino acid sequences each attached to at least one further functional domain;
   wherein any one or more of said amino acid sequences interacts with at least one of said amino acid sequences in a complementary fashion to form a parental, native-like tertiary or optionally quaternary structure and wherein said parental, native-like tertiary or optionally quaternary structure is derived from a single parent polypeptide.

2. The multifunctional polypeptide according to claim 1, wherein said single parent polypeptide is ubiquitin, acyl-phosphatase, IL2, calbindin or apomyoglobin.

3. The multifunctional polypeptide according to claim 1 or 2, wherein said parental, native-like tertiary or optionally quaternary structure is biologically active.

4. The multifunctional polypeptide according to any one of claims 1 to 3, wherein at least one of said functional domains is a fragment derived from a member of the immunoglobulin superfamily.

5. The multifunctional polypeptide according to claim 4, wherein two of said functional domains are fragments derived from members of the immunoglobulin superfamily.

6. The multifunctional polypeptide according to claim 4 or 5, wherein said fragments are antibody fragments.

7. The multifunctional polypeptide according to any one of claims 1 to 6, wherein at least one of said functional domains is a biologically active molecule or a derivative thereof other than a fragment derived from a member of the immunoglobulin superfamily.

8. The multifunctional polypeptide according to any one of claims 1 to 6, wherein the folding of the amino acid sequences is stabilised by a covalent bonding.

9. The multifunctional polypeptide according to any one of claims 1 to 8, wherein at least one of said functional domains is coupled to said amino acid sequence(s) via a flexible peptide linker.

10. The multifunctional polypeptide according to claim 9, wherein said flexible peptide linker is an antibody hinge region.

11. The multifunctional polypeptide according to any one of calims 1 to 10, wherein at least one of said amino acid sequences is coupled to at least one further (poly)peptide.

12. The multifunctional polypeptide according to claim 11, wherein said further (poly)peptide is an enzyme, a toxin, a cytokine, a metal binding site, a metal binding protein, a soluble receptor, a DNA-binding domain, a transcription factor, an immunoglobulin, a bioactive peptide of 5 to 15 amino acid residues, a peptide hormone, a growth factor, a lectin, a lipoprotein, and a peptide which is able to bind to an independent binding entity.

13. The multifunctional polypeptide according to any one of claims 1 to 12, wherein at least one of said amino acid sequences, functional domains or further (poly)peptide(s) is of human origin.

14. A DNA sequence encoding an amino acid sequence and at least one functional domain and, optionally, at least one further functional (poly)peptide comprised in the multifunctional polypeptide of any one of claims 1 to 13.

15. A vector comprising at least one DNA molecule of claim 14.

16. The vector of claim 15, which is a bicistronic vector.

17. A vector cassette characterised in that it comprises a DNA sequence encoding an amino acid sequence and optionally at least one further (poly)peptide comprised in the multifunctional polypeptide of any one of claims 1 to 13, and additionally at least one, preferably a singular cloning site for inserting the DNA encoding at least one further functional domain.

18. A vector cassette characterised in that it comprises DNA sequences encoding the amino acid sequences, and optionally the further (poly)peptide(s) comprised in the multifunctional polypeptide of any one of claims 1 to 13, and suitable restriction sites for the cloning of DNA sequences encoding the functional domains, such that upon expression of the DNA sequences after the insertion of the DNA sequences encoding the functional domains into said restriction sites, in a suitable host the multifunctional polypeptide according to any one of claims 1 to 13 is formed.

19. The vector cassette according to claim 17 or 18 characterised in that it comprises the inserted DNA sequence(s) encoding said functional domain(s).

20. A host cell transformed with at least one vector according to claim 15 or 16, or at least one vector cassette according to claim 19.

21. The host cell according to claim 20, which is a mammalian, perferably human, yeast, insect, plant or bacterial, preferably E. coli cell.

22. A method for the production of a multifunctional polypeptide according to any one of claims 1 to 13, which comprises culturing the host cell according to claim 20 or 21 in a suitable medium, and recovering said multifunctional polypeptide produced by said host cell.

23. A method for the production of a multifunctional polypeptide according to any one of claims 1 to 13 which comprises culturing at least two host cells according to claim 20 or 21 in a suitable medium, said host cells each producing only one of said first and said second amino acid sequences attached to at least one further functional domain, recovering the amino acid sequences, mixing thereof under mildly denaturing conditions and allowing in vitro folding of the multifunctional polypeptide according to any one of claims 1 to 13 from said amino acid sequences.

24. The method according to claim 23, wherein the further amino acid sequence(s) (each) attached to at least one further functional domain are/is produced by at least one further host cell not producing said first or second amino acid sequence.

25. The method according to claim 23, wherein at least one further amino acid sequence attached to at least one further functional domain is produced by the host cell according to claim 20 or 21 producing said first or second amino acid sequence.

26. A pharmaceutical composition comprising the multifunctional polypeptide according to any one of claims 1 to 13 optionally in combination with a pharmaceutically acceptable carrier.

27. A diagnostic composition comprising the multifunctional polypeptide according to any one of claims 1 to 13.

28. A kit comprising at least one vector cassette according to claim 17 or 18.

Fig.1: segmented tertiary structure for a targeted hetero-association

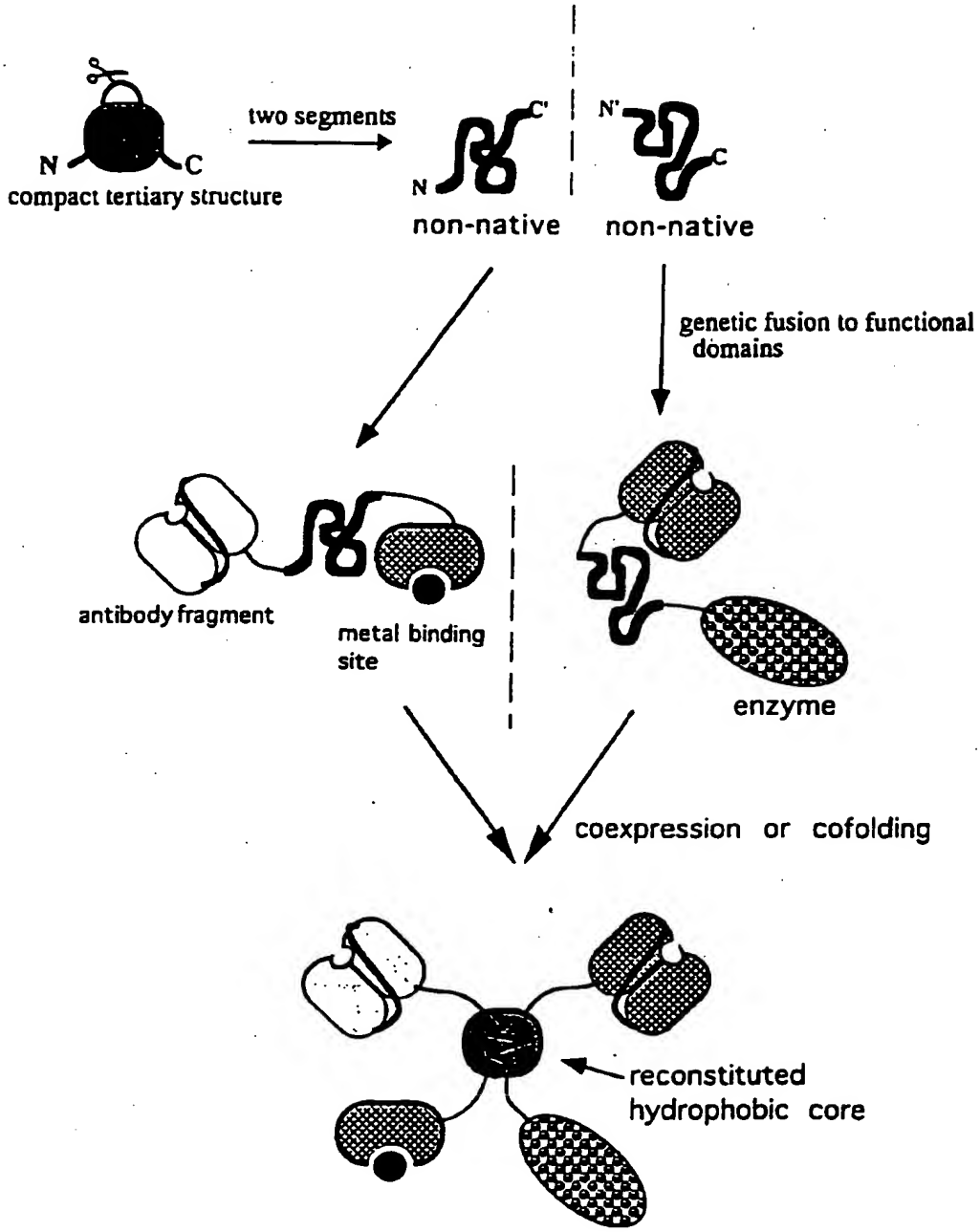


Fig. 2 : Modular cistron encoding functional domains N- and/or C-terminally fused to the assembly device

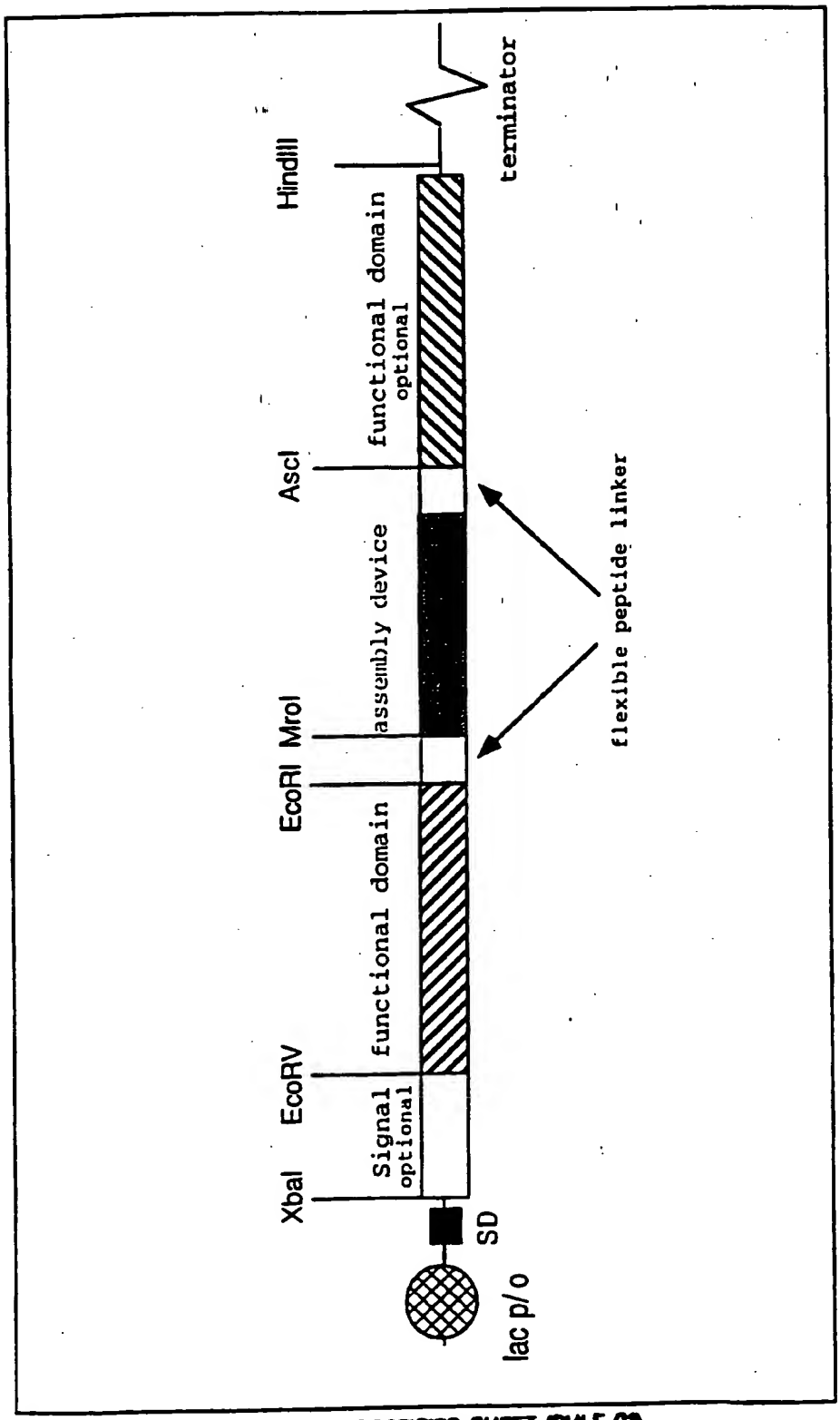


Fig. 3 : protein sequence of human ubiquitin (segmented after Gly35)

```
1         10          20          30      ✂ 40          50
MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL

          60          70
EDGRTLSDYN IQKESTLHLV LRLRGG**
```

Fig. 4: MroI-Hind III gene cassette encoding for segment ALPHA of ubiquitin

```
MroI
 S   G  M   Q   I   F   V   K   T   L   T   G   K   T   I   T   L   E
TCC GGA ATG CAG ATC TTC GTT AAA ACC CTG ACC GGT AAA ACC ATC ACC CTG GAA
           9          18          27          36          45          54
AGG CCT TAC GTC TAG AAG CAA TTT TGG GAC TGG CCA TTT TGG TAG TGG GAC CTT

 V   E   P   S   D   T   I   E   N   V   K   A   K   I   Q   D   K   E
GTT GAA CCG TCT GAC ACC ATC GAA AAC GTT AAA GCT AAA ATC CAG GAC AAA GAA
           63          72          81          90          99         108
CAA CTT GGC AGA CTG TGG TAG CTT TTG CAA TTT CGA TTT TAG GTC CTG TTT CTT

          HindIII
 G   *   *   A
GGT TGA TAA GCT T 3'
       117
CCA ACT ATT CGA A 5'
```

Fig. 5: MroI-Hind III gene cassette encoding for segment BETA of ubiquitin

```
MroI
S   G   I   P   P   D   Q   Q   R   L   I   F   A   G   R   Q   L   E
TCC GGA ATC CCG CCG GAC CAG CAG CGT CTG ATC TTC GCT GGT CGT CAG CTG GAA
          9           18          27          36          45          54
AGG CCT TAG GGC GGC CTG GTC GTC GCA GAC TAG AAG CGA CCA GCA GTC GAC CTT

D   G   R   T   L   S   D   Y   N   I   Q   K   E   S   T   L   H   L
GAC GGT CGT ACC CTG TCT GAC TAC AAC ATC CAG AAA GAA TCT ACC CTG CAC CTG
          63          72          81          90          99          108
CTG CCA GCA TGG GAC AGA CTG ATG TTG TAG GTC TTT CTT AGA TGG GAC GTG GAC

                    HindIII
V   L   R   L   *   *
GTT CTG CGT CTG TGA TAA  3'
        117         126
CAA GAC GCA GAC ACT ATT  5'
```

Fig. 6: EcoRI-MroI gene cassette encoding a flexible linker (huIgG3)

```
    EcoRI                                                MroI
      E   F   T   P   L   G   D   T   T   H   T   S   G
5'  GAA TTC ACC CCG CTG GGT GAC ACC ACC CAC ACC TCC GGA  3'
            9           18          27          36
3'  CTT AAG TGG GGC GAC CCA CTG TGG TGG GTG TGG AGG CCT  5'
```

Fig. 7 : Construction of monocistronic expression vector

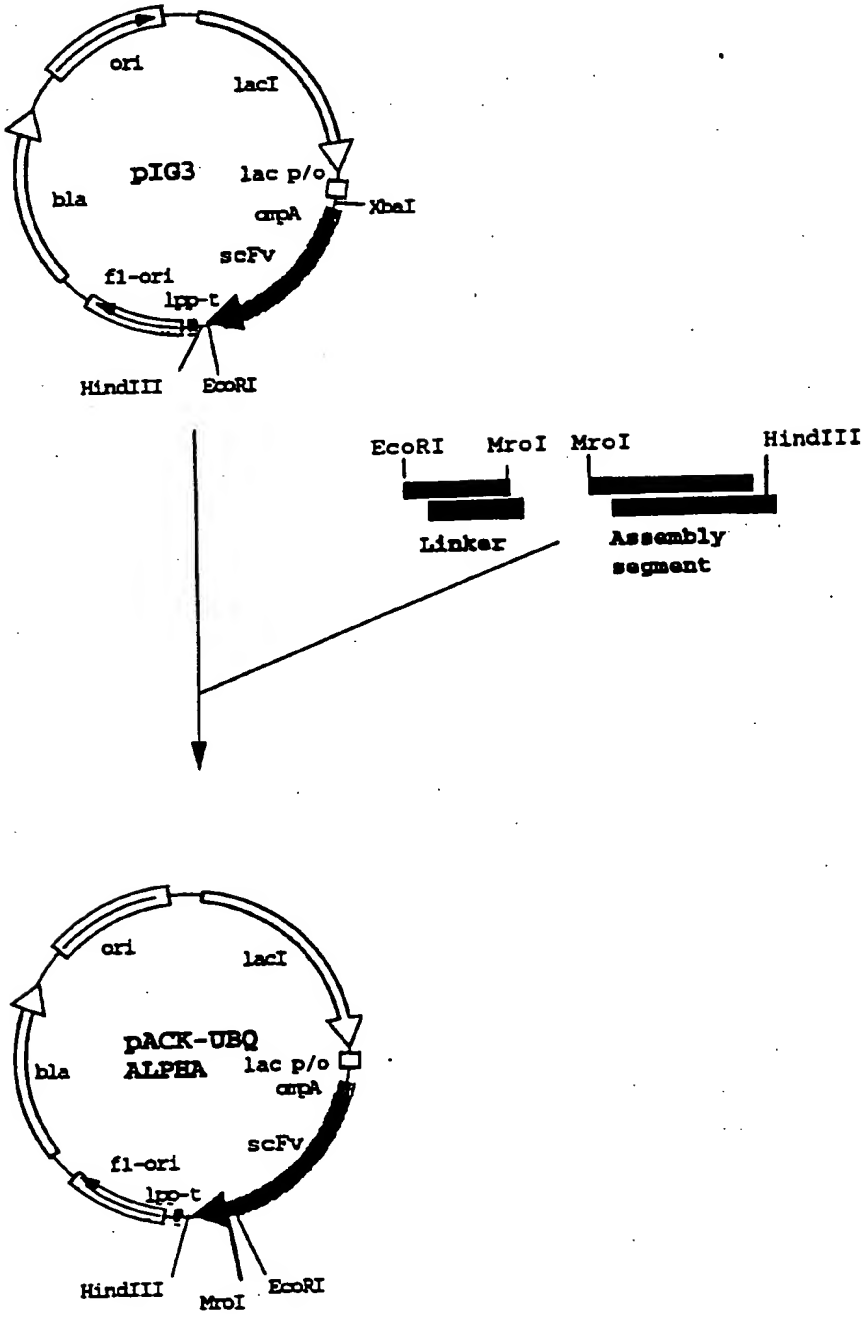
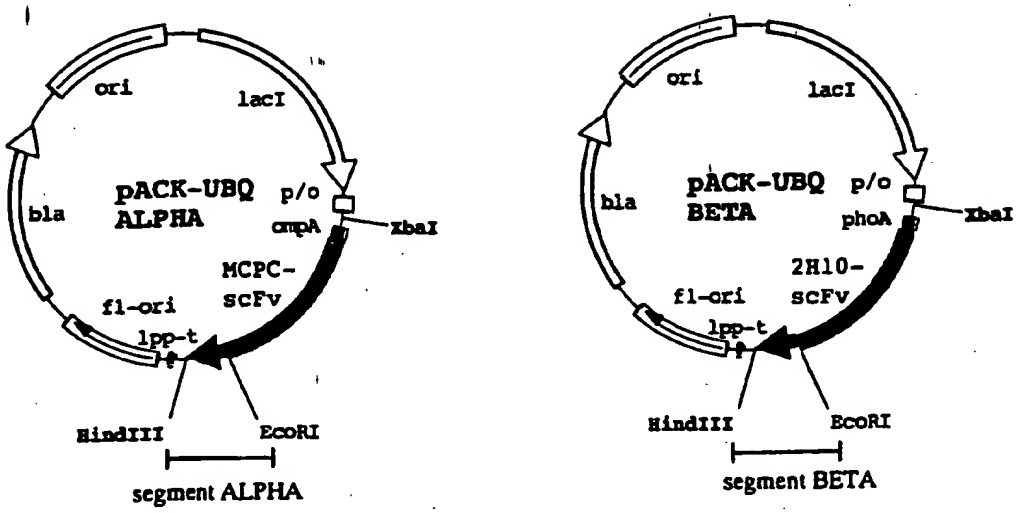


Fig. 8: Construction of dicistronic co-expression vectors



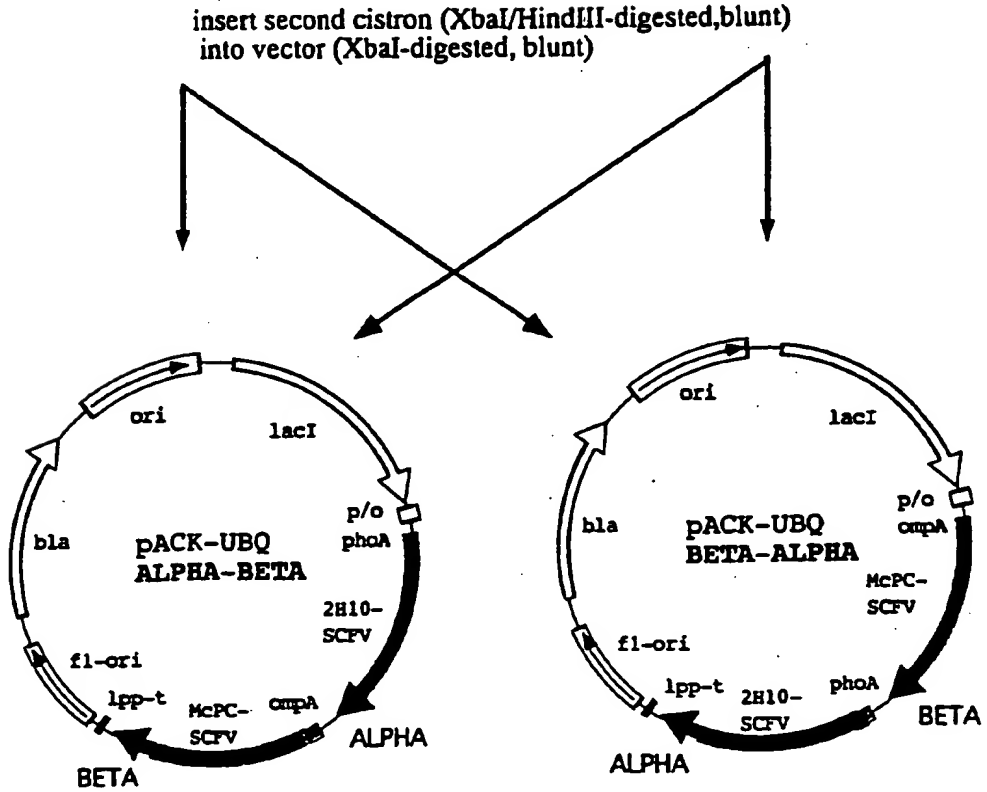


Fig. 9 : protein sequence of human ubiquitin with intersegmental disufides Cys4 and Cys66 (segmented after Gly35)

```
1         10          20          30      ✂ 40          50
MQICVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL

          60          70
EDGRTLSDYN IQKESCLHLV LRLRGG**
```

Fig. 10: MroI-Hind III gene cassette encoding for segment ALPHA-CYS4 of ubiquitin

```
MroI
 S   G   M   Q   I   C   V   K   T   L   T   G   K   T   I   T   L   E
 TCC GGA ATG CAG ATC TGC GTT AAA ACC CTG ACC GGT AAA ACC ATC ACC CTG GAA
           9           18          27          36          45          54
 AGG CCT TAC GTC TAG ACG CAA TTT TGG GAC TGG CCA TTT TGG TAG TGG GAC CTT

 V   E   P   S   D   T   I   E   N   V   K   A   K   I   Q   D   K   E
 GTT GAA CCG TCT GAC ACC ATC GAA AAC GTT AAA GCT AAA ATC CAG GAC AAA GAA
           63          72          81          90          99          108
 CAA CTT GGC AGA CTG TGG TAG CTT TTG CAA TTT CGA TTT TAG GTC CTG TTT CTT

          HindIII
 G   *   *   A
 GGT TGA TAA GCT T 3'
       117
 CCA ACT ATT CGA A 5'
```

Fig. 11 : MroI-AscI-Hind III gene cassette encoding for segment BETA-CYS66 with C-terminal GGSGGAP linker of ubiquitin

```
MroI
S   G   I   P   P   D   Q   Q   R   L   I   F   A   G   R   Q   L   E
TCC GGA ATC CCG CCG GAC CAG CAG CGT CTG ATC TTC GCT GGT CGT CAG CTG GAA
          9           18          27          36          45          54
AGG CCT TAG GGC GGC CTG GTC GTC GCA GAC TAG AAG CGA CCA GCA GTC GAC CTT

D   G   R   T   L   S   D   Y   N   I   Q   K   E   S   C   L   H   L
GAC GGT CGT ACC CTG TCT GAC TAC AAC ATC CAG AAA GAA TCT TGC CTG CAC CTG
          63          72          81          90          99          108
CTG CCA GCA TGG GAC AGA CTG ATG TTG TAG GTC TTT CTT AGA ACG GAC GTG GAC

                                        AscI            HindIII
V   L   R   L   G   G   S   G   G   A   P   *   *
GTT CTG CGT CTG GGG GGG AGC GGA GGC GCG CCG TGA TAA   3'
        117         126
CAA GAC GCA GAC CCC CCC TCG CCT CCG CGC GGC ACT ATT   5'
```

Fig. 12: Protein sequence of human IL-2 (segmented after His79)

```
        10          20          30          40          50          60
APTSSSTKKT QLQLEHLLLD LQMILNGINN YKNPKLTRML TFKFYMPKKA TELKHLQCLE

        70          ✂           90          100         110         120
EELKPLEEVL NLAQSKNFHL RPRDLISNIN VIVLELKGSE TTFMCEYADE TATIVEFLNR

       130
WITFCQSIIS TLT
```

Fig.13 : MroI-AscI-Hind III gene cassette encoding for segment ALPHA of human IL-2

```
MroI
S   G   A   P   T   S   S   S   T   K   K   T   Q   L   Q   L   E   H
TCC GGA GCA CCT ACT TCA AGT TCT ACA AAG AAA ACA CAG CTA CAA CTG GAG CAT
          9           18          27          36          45          54
AGG CCT CGT GGA TGA AGT TCA AGA TGT TTC TTT TGT GTC GAT GTT GAC CTC GTA

L   L   L   D   L   Q   M   I   L   N   G   I   N   N   Y   K   N   P
TTA CTG CTG GAT TTA CAG ATG ATT TTG AAT GGA ATT AAT AAT TAC AAG AAT CCC
          63          72          81          90          99          108
AAT GAC GAC CTA AAT GTC TAC TAA AAC TTA CCT TAA TTA TTA ATG TTC TTA GGG

K   L   T   R   M   L   T   F   K   F   Y   M   P   K   K   A   T   E
AAA CTC ACC AGG ATG CTC ACA TTT AAG TTT TAC ATG CCC AAG AAG GCC ACA GAA
          117         126         135         144         153         162
TTT GAG TGG TCC TAC GAG TGT AAA TTC AAA ATG TAC GGG TTC TTC CGG TGT CTT

L   K   H   L   Q   C   L   E   E   E   L   K   P   L   E   E   V   L
CTG AAA CAT CTT CAG TGT CTA GAA GAA GAA CTC AAA CCT CTG GAG GAA GTG CTA
          171         180         189         198         207         216
GAC TTT GTA GAA GTC ACA GAT CTT CTT CTT GAG TTT GGA GAC CTC CTT CAC GAT

                                                        AscI    HindIII
N   L   A   Q   S   K   N   F   H   G   G   S   G   G   A   P   *
AAT TTA GCT CAA AGC AAA AAC TTT CAC GGG GGG AGC GGA GGC GCG CCG TGA T
          225         234         243         252         261
TTA AAT CGA GTT TCG TTT TTG AAA GTG CCC CCC TCG CCT CCG CGC GGC ACT A
```

Fig. 14 : MroI-AscI-Hind III gene cassette encoding for segment BETA of human IL-2

```
MroI
S   G   L   R   P   R   D   L   I   S   N   I   N   V   I   V   L   E
TCC GGA TTA AGA CCC AGG GAC TTA ATC AGC AAT ATC AAC GTA ATA GTT CTG GAA
          9           18          27          36          45          54
AGG CCT AAT TCT GGG TCC CTG AAT TAG TCG TTA TAG TTG CAT TAT CAA GAC CTT

L   K   G   S   E   T   T   F   M   C   E   Y   A   D   E   T   A   T
CTA AAG GGA TCT GAA ACA ACA TTC ATG TGT GAA TAT GCT GAT GAG ACA GCA ACC
          63          72          81          90          99          108
GAT TTC CCT AGA CTT TGT TGT AAG TAC ACA CTT ATA CGA CTA CTC TGT CGT TGG

I   V   E   F   L   N   R   W   I   T   F   C   Q   S   I   I   S   T
ATT GTA GAA TTT CTG AAC AGA TGG ATT ACC TTT TGT CAA AGC ATC ATC TCA ACA
          117         126         135         144         153         162
TAA CAT CTT AAA GAC TTG TCT ACC TAA TGG AAA ACA GTT TCG TAG TAG AGT TGT

                                AscI        HindIII
L   T   G   G   S   G   G   A   P   *
CTG ACT GGG GGG AGC GGA GGC GCG CCG TGA T 3'
          171         180         189
GAC TGA CCC CCC TCG CCT CCG CGC GGC ACT A 5'
```

Fig. 15 Protein sequence of human apomyoglobin (cut after Lys47 and Lys98)

mglsdgewql vlnvwgkvea dipghgqevl irlfkghpet lekfdkfkhl

51
ksedemkase dlkkhgatvl talggilkkk ghheaeikpl aqshatkhki

101
pvkylefise ciiqvlqskh pgdfgadaeg amnkalelfr kdmasnykel

151
gfqg

Fig. 16 : MroI-AscI-Hind III gene cassette encoding for segment ALPHA of human apomyoglobin

```
MroI
S   G   M   G   L   S   D   G   E   W   Q   L   V   L   N   V   W   G
TCC GGA ATG GGT CTG TCT GAC GGT GAA TGG CAG CTG GTT CTG AAC GTT TGG GGT
          9          18          27          36          45          54
AGG CCT TAC CCA GAC AGA CTG CCA CTT ACC GTC GAC CAA GAC TTG CAA ACC CCA


K   V   E   A   D   I   P   G   H   G   Q   E   V   L   I   R   L   F
AAA GTT GAA GCT GAC ATC CCG GGT CAC GGT CAG GAA GTT CTG ATC CGT CTG TTC
          63          72          81          90          99         108
TTT CAA CTT CGA CTG TAG GGC CCA GTG CCA GTC CTT CAA GAC TAG GCA GAC AAG


K   G   H   P   E   T   L   E   K   F   D   K   F   K   G   G   S   G
AAA GGT CAC CCG GAA ACC CTG GAA AAA TTC GAC AAA TTC AAA GGG GGG AGC GGA
        117         126         135         144         153         162
TTT CCA GTG GGC CTT TGG GAC CTT TTT AAG CTG TTT AAG TTT CCC CCC TCG CCT


AscI          HindIII
 G   A   P    *
GGC GCG CCG TGA T 3'
        171
CCG CGC GGC ACT A 5'
```

Fig. 17: MroI-AscI-Hind III gene cassette encoding for segment BETA of human apomyoglobin

```
MroI
S   G   H   L   K   S   E   D   E   M   K   A   S   E   D   L   K   K
TCC GGA CAC CTG AAA TCT GAA GAC GAA ATG AAA GCA TCT GAA GAC CTG AAA AAA
          9           18          27          36          45          54
AGG CCT GTG GAC TTT AGA CTT CTG CTT TAC TTT CGT AGA CTT CTG GAC TTT TTT

H   G   A   T   V   L   T   A   L   G   G   I   L   K   K   K   G   H
CAC GGT GCT ACC GTT CTG ACC GCT CTG GGT GGT ATC CTG AAA AAA AAA GGT CAC
          63          72          81          90          99          108
GTG CCA CGA TGG CAA GAC TGG CGA GAC CCA CCA TAG GAC TTT TTT TTT CCA GTG

H   E   A   E   I   K   P   L   A   Q   S   H   A   T   K   H   K   G
CAC GAA GCT GAA ATC AAA CCG CTG GCT CAG TCT CAC GCT ACC AAA CAC AAA GGG
          117         126         135         144         153         162
GTG CTT CGA CTT TAG TTT GGC GAC CGA GTC AGA GTG CGA TGG TTT GTG TTT CCC

                AscI            HindIII
 G   S   G   G   A   P   *
GGG AGC GGA GGC GCG CCG TGA T 3'
          171         180
CCC TCG CCT CCG CGC GGC ACT A 5'
```

Fig.18: MroI-AscI-Hind III gene cassette encoding for segment GAMMA of human apomyoglobin

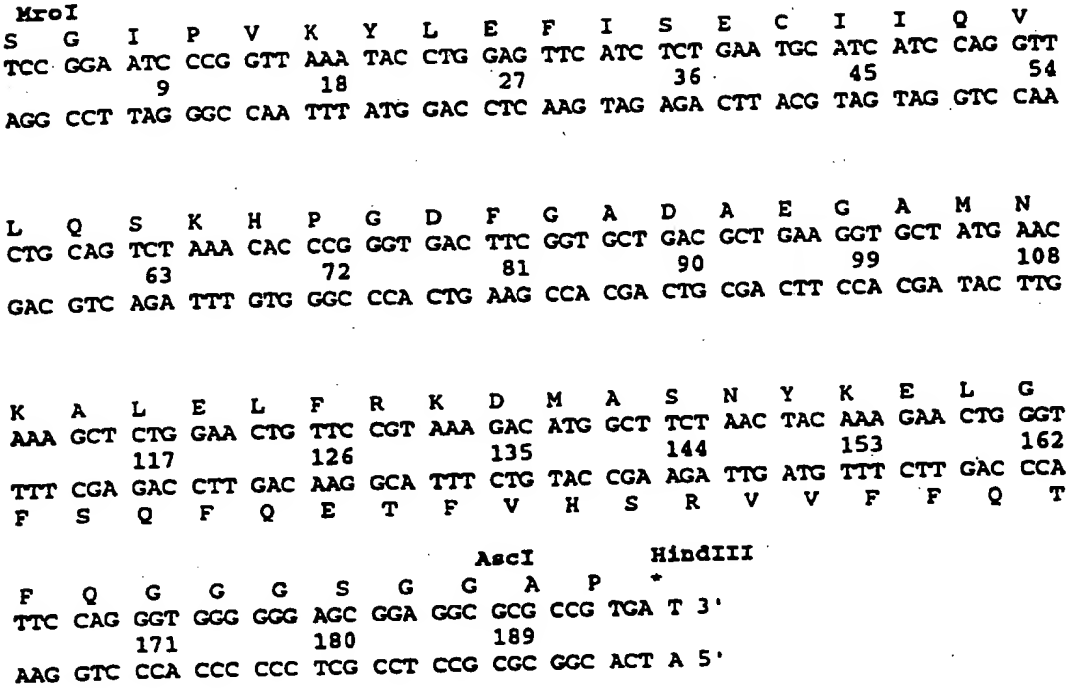


Fig. 19: Peptide sequence of an endotoxin-neutralizing peptide as a functional domain

```
1          11
RWKVRKSFFKL Q
```

Fig.20: N-terminal EcoRV-EcoRI cassette encoding an endotoxin-neutralizing peptide

```
   EcoRV                                                    EcoRI
    I   M   R   W   K   V   R   K   S   F   F   K   L   Q   E   F
5' ATC ATG CGT TGG AAA GTT CGT AAA TCT TTC TTC AAA CTG CAG GAA TTC  3'
             9           18          27          36          45
3' TAG TAC GCA ACC TTT CAA GCA TTT AGA AAG AAG TTT GAC GTC CTT AAG  5'
```

Fig.21: C-terminal AscI-HindIII cassette encoding an endotoxin-neutralizing peptide

```
   AscI                                                     HindIII
    A   P   R   W   K   V   R   K   S   F   F   K   L   Q   *   *
5' GCG CCG CGT TGG AAA GTT CGT AAA TCT TTC TTC AAA CTG CAG TGA TAA  3'
             9           18          27          36          45
3' CGC GGC GCA ACC TTT CAA GCA TTT AGA AAG AAG TTT GAC GTC ACT ATT  5'
```

Fig. 22 AscI-HINDIII Gene cassette encoding a purification tail for IMAC

```
   AscI                                  Hind III
    A   P   H   H   H   H   H   H   *   *
5' GCG CCG CAC CAC CAC CAC CAC CAC TGA TAA  3'
             9           18          27
3' CGC GGC GTG GTG GTG GTG GTG CAC ACT ATT  5'
```

Fig. 23 Protein sequence of the platelet aggregation inhibitor decorsin as a functional domain

```
1          11          21          31
APRLPQCQGD DQEKCLCNKD ECPPGQCRFP RGDADPYCE
```

Fig. 24 N-terminal EcoRV-EcoRI cassette encoding the platelet aggregation inhibitor decorsin

```
EcoRV
D   I   A   P   R   L   P   Q   C   Q   G   D   D   Q   E   K   C   L
GAT ATC GCT CCG CGT CTG CCG CAG TGC CAG GGT GAC GAC CAG GAA AAA TGC CTG
          9           18          27          36          45          54
CTA TAG CGA GGC GCA GAC GGC GTC ACG GTC CCA CTG CTG GTC CTT TTT ACG GAC

C   N   K   D   E   C   P   P   G   Q   C   R   F   P   R   G   D   A
TGC AAC AAA GAC GAA TGC CCG CCG GGT CAG TGC CGT TTC CCG CGT GGT GAC GCT
          63          72          81          90          99          108
ACG TTG TTT CTG CTT ACG GGC GGC CCA GTC ACG GCA AAG GGC GCA CCA CTG CGA

                EcoRI
D   P   Y   C   E   F
GAC CCG TAC TGC GAA TTC  3'
        117         126
CTG GGC ATG ACG CTT AAG  5'
```

Fig. 25 C-terminal AscI-HindIII cassette encoding the platelet aggregation inhibitor decorsin

```
 AscI
A   P   A   P   R   L   P   Q   C   Q   G   D   D   Q   E   K   C   L
GCG CCG GCT CCG CGT CTG CCG CAG TGC CAG GGT GAC GAC CAG GAA AAA TGC CTG
          12          21          30          39          48          57
CGC GGC CGA GGC GCA GAC GGC GTC ACG GTC CCA CTG CTG GTC CTT TTT ACG GAC

C   N   K   D   E   C   P   P   G   Q   C   R   F   P   R   G   D   A
TGC AAC AAA GAC GAA TGC CCG CCG GGT CAG TGC CGT TTC CCG CGT GGT GAC GCT
          66          75          84          93          102         111
ACG TTG TTT CTG CTT ACG GGC GGC CCA GTC ACG GCA AAG GGC GCA CCA CTG CGA

                  HindIII
D   P   Y   C   E   *   *
GAC CCG TAC TGC GAA TGA TAA  3'
        120         129
CTG GGC ATG ACG CTT ACT ATT  5'
```